

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: MOORE, PAUL A.  
RUBEN, STEVEN M.  
EBNER, REINHARD

(ii) TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE

(iii) NUMBER OF SEQUENCES: 16

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
(B) STREET: 9410 KEY WEST AVENUE  
(C) CITY: ROCKVILLE  
(D) STATE: MD  
(E) COUNTRY: USA  
(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/084,491  
(B) FILING DATE: 27-MAY-1998  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: BROOKES, ANDERS A.  
(B) REGISTRATION NUMBER: 36,373  
(C) REFERENCE/DOCKET NUMBER: PF378

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (301) 309-8504  
(B) TELEFAX: (301) 309-8439

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2329 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 124..913

(ix) FEATURE:

(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 124..184

(ix) FEATURE:

(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 187..913

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTACCAGAAC AGCATAACAA GGGCAGGTCT GACTGCAAGC TGGGACTGGG AGGCAGAGCC 60

GCCGCCAAGG GGGCCTCGGT TAAACACTGG TCGTTCAATC ACCTGCAAGA CGAAGAGGCA 120

AGG ATG CTG TTG GCC TGG GTA CAA GCA TTC CTC GTC AGC AAC ATG CTC 168  
Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu  
-21 -20 -15 -10

CTA GCA GAA GCC TAT GGA TCT GGA GGC TGT TTC TGG GAC AAC GGC CAC 216  
Leu Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His  
-5 1 5 10

CTG TAC CGG GAG GAC CAG ACC TCC CCC GCG CCG GGC CTC CGC TGC CTC 264  
Leu Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu  
15 20 25

AAC TGG CTG GAC GCG CAG AGC GGG CTG GCC TCG GCC CCC GTG TCG GGG 312  
Asn Trp Leu Asp Ala Gln Ser Gly Leu Ala Ser Ala Pro Val Ser Gly  
30 35 40

GCC GGC AAT CAC AGT TAC TGC CGA AAC CCG GAC GAG GAC CCG CGC GGG 360  
Ala Gly Asn His Ser Tyr Cys Arg Asn Pro Asp Glu Asp Pro Arg Gly  
45 50 55

CCC TGG TGC TAC GTC AGT GGC GAG GCC GGC GTC CCT GAG AAA CGG CCT 408  
Pro Trp Cys Tyr Val Ser Gly Glu Ala Gly Val Pro Glu Lys Arg Pro  
60 65 70

TGC GAG GAC CTG CGC TGT CCA GAG ACC ACC TCC CAG GCC CTG CCA GCC 456  
Cys Glu Asp Leu Arg Cys Pro Glu Thr Thr Ser Gln Ala Leu Pro Ala  
75 80 85 90

TTC ACG ACA GAA ATC CAG GAA GCG TCT GAA GGG CCA GGT GCA GAT GAG 504  
Phe Thr Thr Glu Ile Gln Glu Ala Ser Glu Gly Pro Gly Ala Asp Glu  
95 100 105

GTG CAG GTG TTC GCT CCT GCC AAC GCC CTG CCC GCT CGG AGT GAG GCG 552  
Val Gln Val Phe Ala Pro Ala Asn Ala Leu Pro Ala Arg Ser Glu Ala  
110 115 120

GCA GCT GTG CAG CCA GTG ATT GGG ATC AGC CAG CGG GTG CGG ATG AAC 600  
Ala Ala Val Gln Pro Val Ile Gly Ile Ser Gln Arg Val Arg Met Asn  
125 130 135

TCC AAG GAG AAA AAG GAC CTG GGA ACT CTG GGC TAC GTG CTG GGC ATT 648  
Ser Lys Glu Lys Lys Asp Leu Gly Thr Leu Gly Tyr Val Leu Gly Ile  
140 145 150

ACC ATG ATG GTG ATC ATC ATT GCC ATC GGA GCT GGC ATC ATC TTG GGC 696  
Thr Met Met Val Ile Ile Ile Ala Ile Gly Ala Gly Ile Ile Leu Gly  
155 160 165 170

TAC TCC TAC AAG AGG GGG AAG GAT TTG AAA GAA CAG CAT GAT CAG AAA 744  
Tyr Ser Tyr Lys Arg Gly Lys Asp Leu Lys Glu Gln His Asp Gln Lys  
175 180 185

al  
cont

GTA TGT GAG AGG GAG ATG CAG CGA ATC ACT CTG CCC TTG TCT GCC TTC	792
Val Cys Glu Arg Glu Met Gln Arg Ile Thr Leu Pro Leu Ser Ala Phe	
190 195 200	
ACC AAC CCC ACC TGT GAG ATT GTG GAT GAG AAG ACT GTC GTG GTC CAC	840
Thr Asn Pro Thr Cys Glu Ile Val Asp Glu Lys Thr Val Val Val His	
205 210 215	
ACC AGC CAG ACT CCA GTT GAC CCT CAG GAG GGC AGC ACC CCC CTT ATG	888
Thr Ser Gln Thr Pro Val Asp Pro Gln Glu Gly Ser Thr Pro Leu Met	
220 225 230	
GGC CAG GCC GGG ACT CCT GGG GCC T GAGCCCCCCC AGTGGGCAGG	933
Gly Gln Ala Gly Thr Pro Gly Ala	
235 240	
AGCCCATGCA GACACTGGTG CAGGACAGCC CACCCTCCTA CAGCTAGGAG GAACTACCAC	993
TTTGTGTTCT GGTTAAAACC CTACCACTCC CCCGCTTTTT TGGCGAATCC TAGTAAGAGT	1053
GACAGAAGCA GGTGGCCCTG TGGGCTGAGG GTAAGGCTGG GTAGGGTCCT AACAGTGCTC	1113
CTTGTCCATC CCTTGGAGCA GATTTTGTCT GTGGATGGAG ACAGTGGCAG CTCCCACAGT	1173
GATGCTGCTG CTAAGGGCTT CCAAACATTG CCTGCACCCC TGGAAGTAA CCAGGGATAG	1233
ACGGGGAGCT CCCCCAGGCT CCTCTGTGCT TTAATAAGAT GGCTCAGTCT CCACTGTGGG	1293
CTTGAGTGGC ATACACTGTT ATTCATGGTT AAGGTAAAGC AGGTCAAGGG ATGGCATTGA	1353
AAAAATATAT TTAGTTTTTA AAATATTTGG GATGGAACTC CCTACTGACC TCTGACAACT	1413
GGAAACGAGT TTGTACTGAA GTCAGAACTT TGGGTTGGGA ATGAGATCTA GGTTGTGGCT	1473
GCTGGTATGC TTCAGCTTGC TGGCAATGAT GTGCCTTGAC AACCGTGGGC CAGGCCTGGG	1533
CCCAGGGACT CTTCTGTGTT CATAAGGAAA GGAAGAATTG CACTGAGCAT TCCACTTAGG	1593
AAGAGGATAG AGAAGGATCT GCTCCGCCTT TGGCCACAGG AGCAGAGGCA GACCTGGGAT	1653
GCCCCAGTTT CTCTTCAGGG ATGGATAGTG ACCTGTCTTC ATTTTGCACA GGTAAGAGAG	1713
TAGTTAGCTA ACCTATGGGA ATTATACTGT GGGGCCTTGT GAGCTGCTTC TAAGAGGCTA	1773
ACCTGGAAAC TAAGCTCAGA GGCAAGGTAA TAAAGCACTT CAGGGCTTGC TCCCCAAGTG	1833
GGCCTGATTT AGCAGGTGGT CTGCGGGCGT CCAGGTCAGC ACCTTCCTGT AGGGCACTGG	1893
GGCTAGGGTC ACAGCCCCCTA ACTCATAAAG CAATCAAAGA ACCATTAGAA AGGGCTCATT	1953
AAGCCTTTTG GACACAGGAC CCCAGAGAGG AAAAAGTGAC TTGCCCAAGG TCGTAAGCAA	2013
GCTACTGGCA TGGCAAGAGC CCAGCTTCCT GACGGAGCGC AACATTTCTC CACTGCACTG	2073
TGCTAGCAGC TCAGCAGGGC CTCTAACCTG TGATGTCACA CTCAAGAGGC CTTGGCAGCT	2133
CCTAGCCATA GAGCTTCCTT TCCAGAACCC TTCCACTGCC CAATGTGGAG ACAGGGGTTA	2193
GTGGGGCTTT CTATGGAGCC ATCTGCTTTG GGGACCTAGA CCTCAGGTGG TCTCTTGGTG	2253

a!  
cont

TTAGTGATGC TGGAGAAGAG AATATTACTG GTTCTACTT TTCTATAAAG GCATTTCTCT 2313  
 ATAAAAAAAA AAAAAA 2329

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 263 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu Leu  
 -21 -20 -15 -10

Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His Leu  
 -5 1 5 10

Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu Asn  
 15 20 25

Trp Leu Asp Ala Gln Ser Gly Leu Ala Ser Ala Pro Val Ser Gly Ala  
 30 35 40

Gly Asn His Ser Tyr Cys Arg Asn Pro Asp Glu Asp Pro Arg Gly Pro  
 45 50 55

Trp Cys Tyr Val Ser Gly Glu Ala Gly Val Pro Glu Lys Arg Pro Cys  
 60 65 70 75

Glu Asp Leu Arg Cys Pro Glu Thr Thr Ser Gln Ala Leu Pro Ala Phe  
 80 85 90

Thr Thr Glu Ile Gln Glu Ala Ser Glu Gly Pro Gly Ala Asp Glu Val  
 95 100 105

Gln Val Phe Ala Pro Ala Asn Ala Leu Pro Ala Arg Ser Glu Ala Ala  
 110 115 120

Ala Val Gln Pro Val Ile Gly Ile Ser Gln Arg Val Arg Met Asn Ser  
 125 130 135

Lys Glu Lys Lys Asp Leu Gly Thr Leu Gly Tyr Val Leu Gly Ile Thr  
 140 145 150 155

Met Met Val Ile Ile Ala Ile Gly Ala Gly Ile Ile Leu Gly Tyr  
 160 165 170

Ser Tyr Lys Arg Gly Lys Asp Leu Lys Glu Gln His Asp Gln Lys Val  
 175 180 185

Cys Glu Arg Glu Met Gln Arg Ile Thr Leu Pro Leu Ser Ala Phe Thr  
 190 195 200

Asn Pro Thr Cys Glu Ile Val Asp Glu Lys Thr Val Val Val His Thr  
 205 210 215

1  
cont.

Ser Gln Thr Pro Val Asp Pro Gln Glu Gly Ser Thr Pro Leu Met Gly  
 220 225 230 235

Gln Ala Gly Thr Pro Gly Ala  
 240

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Tyr Val Phe Lys Ala Gly Lys Tyr Ser Ser Glu Phe Cys Ser Thr Pro  
 1 5 10 15

Ala Cys Ser Glu Gly Asn Ser Asp Cys Tyr Phe Gly Asn Gly Ser Ala  
 20 25 30

Tyr Arg Gly Thr His Ser Leu Thr Glu Ser Gly Ala Ser Cys Leu Pro  
 35 40 45

Trp Asn Ser Met Ile Leu Ile Gly Lys Val Tyr Thr Ala Gln Asn Pro  
 50 55 60

Ser Ala Gln Ala Leu Gly Leu Gly Lys His Asn Tyr Cys Arg Asn Pro  
 65 70 75 80

Asp Gly Asp Ala Lys Pro Trp Cys His Val Leu Lys Asn Arg Arg Leu  
 85 90 95

Thr Trp Glu Tyr Cys Asp Val Pro Ser Cys Ser Thr Cys Gly Leu Arg  
 100 105 110

Gln Tyr Ser Gln Pro Gln Phe Arg Ile Lys Gly Gly Leu Phe Ala Asp  
 115 120 125

Ile Ala Ser His Pro Trp Gln Ala Ala Ile Phe Ala Lys His Arg Arg  
 130 135 140

Ser Pro Gly Glu Arg Phe Leu Cys Gly Gly Ile Leu Ile Ser Ser Cys  
 145 150 155 160

Trp Ile Leu Ser Ala Ala His Cys Phe Gln Glu Arg Phe Pro Pro His  
 165 170 175

His Leu Thr Val Ile Leu Gly Arg Thr Tyr Arg Val Val Pro Gly Glu  
 180 185 190

Glu Glu Gln Lys Phe Glu Val Glu Lys Tyr Ile Val His Lys Glu Phe  
 195 200 205

Asp Asp Asp Thr Tyr Asp Asn Asp Ile Ala Leu Leu Gln Leu Lys Ser

Q1  
 Cont.

210	215	220
Asp Ser Ser Arg Cys Ala Gln Glu Ser Ser Val Val Arg Thr Val Cys		
225	230	235 240
Leu Pro Pro Ala Asp Leu Gln Leu Pro Asp Trp Thr Glu Cys Glu Leu		
	245	250 255
Ser Gly Tyr Gly Lys His Glu Ala Leu Ser Pro Phe Tyr Ser Glu Arg		
	260	265 270
Leu Lys Glu Ala His Val Arg Leu Tyr Pro Ser Ser Arg Cys Thr Ser		
	275	280 285
Gln His Leu Leu Asn Arg Thr Val Thr Asp Asn Met Leu Cys Ala Gly		
	290	295 300
Asp Thr Arg Ser Gly Gly Pro Gln Ala Asn Leu His Asp Ala Cys Gln		
	305	310 315 320
Gly Asp Ser Gly Gly Pro Leu Val Cys Leu Asn Asp Gly Arg Met Thr		
	325	330 335
Leu Val Gly Ile Ile Ser Trp Gly Leu Gly Cys Gly Gln Lys Asp Val		
	340	345 350
Pro Gly Val Tyr Thr Lys Val Thr Asn Tyr Leu Asp Trp Ile Arg Asp		
	355	360 365
Asn Met Arg Pro		
	370	

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 250 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATTGCACTGA GCATTCCACT TAGGAAGAGG ATAGAGAAGG ATCTGCTCCG CCTTTGGCCA 60  
 CAGGAGCAGA GGCAGACCTG GGATGCCCCA TTTCTCTTCA GGGATGGATA GTGACCTGTC 120  
 TTCATTTTGC ACAGGTAAGA GAGTAGTTAG CTAACCTATG GGAATTATAC TGTGGGGCCT 180  
 TGTAGCTGCT TCTAAGAGGC TAACCTGGAA ACTAAGCTCA GAGGCAAGGT AATAAAGCAC 240  
 TTCAGGGCTT 250

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 247 base pairs

Q1  
Cont.

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATAGAGAAAT GCCTTTATAG AAAAGTAGAA ACCAGTAATA TTCTCTTCTC CAGCATCACT	60
AACACCAAGA GACCACCTGA GGTCTAGGTC CCCAAAGCAG ATGGCTCCAT AGAAAGCCCC	120
ACTAACCCGT CTCCACATTG GGCAGTGGAA GGGTTCTGGA AAGGAAGCTC TATGGCTAGG	180
AGCTGCCAAG GCCTCTTGAG TGTGACATCA CAGGTTAGAG GCCCTGCTGA GCTGCTAGCA	240
CAGTGCA	247

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 461 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

*a1 cont*

AATTCGGCAA GAGTAACAGC ATAACAAGGG TAGGTCTGAC TGCAGCTGGG ACTGGGAGGC	60
AGAGCACGCC AAGGGGGCCT CGGTAAACA CTGGTCGTTT AATCACCTGC AAACGAGGAG	120
GCAAGGATGC TGTTGGCCTG GGTACAGCAT TCCTGGTCAG CAACATGCTC CTAGCGTAAG	180
CCTATGGATC TGGAGGCTGT TTCTGGGACA ACGGCCACTG TACCCGGAGG ACCAGACCTT	240
CCCGGCCGGT CCTCGTGCCT CAACTGGCTG GACGCGCAGG GCTGCCTGGG CCCCTTTTTC	300
GGTCAAATTT CACAGTTTAC TTCGAAACCG GGACGGGGCC GTGGGGGCCC TGGTGGTTAG	360
TTTGGGGTCG GGTTTTCTTA AAAAAGGTTT TTGGGGCCGG TTTTCGGAAC CATTTGCGTT	420
GAATTTTTTA GGGAAATTC AGGAGTTTTT TAAGGGCCAT T	461

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGCAAGTTGC AGAACTGGAA ACGAGTTTGT ACAGAAGTCA GAACTTTGGG TTAGGAATGA	60
GATCTAGGTT GTGGCTGCTG GTATGCTTCA TTGCTGGCAA TAATGTGCCT TGACAACCGT	120
GGGCCAGGCC TGGGACCAGG GACTCTTCCT GTTTCATAAG GAAAGGAAGA ATTGCACTGA	180
GCATTCCACT TAGGAAGAGG ATAGAGCAAG GAATCTGCTC CGCTTTGGCC ACAGGAGCAG	240
AGGCAGACCT GGGATGCCCC AGTTCTCTTT CAGGGATGGG ATAGTGACCT GTCTTACATT	300
TTGCACAGGT AAAGAGAGTT AGTTAGCTAA CCTATTGGGC TTTATTACTT GGGGCTTGTG	360
AGCTGCTTTT TAAGAGGTTA ACCTGGAACCT AAAGTTCAG	399

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 334 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

*a1 cont.*

TAATTCGGCA AGGGACAGGT CTGACTGCAG CTGGGACTGG GAGGCAGAGC CGTCAAGGGG	60
GCCTCGGTTA AACACTGGTC GTTCAATCAC CTGCAACGAG AGGCAAGGAT GCTGTTGGCC	120
TGGGTACAAG CATTCTGTC AGCAACATGC TCCTAGCAGA AAGCCTATGG ATCTGGGAGG	180
CTGTTTCTGG GACAACGGCC ACCTGTACCG GAGGACCAGA CCTCCCCGGC CGGGCCTTCC	240
GTGGCCTTCA ATTGGTTTGA CGTGGCAAAG GGGCTTGTCT GGCCCTTTTG GGGGAAAATT	300
ACAAGTTTTA ATTGTCCCGG AAAACCTGGA GAGG	334

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:



AATTCGGCAG AGGGAGAGGG AGATGCAGCG AATCACTCTG CCCTTGTCTG CCTTCACCAA 60  
 CCCCACCTGT GAGATTGTGG ATGAGAAGAC TGTCGTGGTC CACACCAGCC AGACTCCAGT 120  
 TGACCCTCAG GAGGGCAGCA CCCCCTTAT GGGACCAGGC CGGGGACTCC TGGGGCCTGA 180  
 GCCCCCAGT GGGGCAGGAG CCATGGCAGA CACTGGTGCA GGACAGCCAC CCTCCTTACA 240  
 GCTAGGGGGA ACTACCACTT TGTGTTTCTG GTTTAAAACC CTACCACTCC CGGATTTTTT 300  
 GGCGGATTCC TTAGTTAAGA GTACAGAAGC AGGTGGGCCT ATGGCTTGGA GGGTAAGGTG 360  
 GGGTAGGGTT CCTAAAAGTG GGTTCCTGGT TGCTCCTGGG AGGAAGATT TGGTTTTGGT 420  
 GGGGACAGTG GCAGTTTCCA CAGGTTGTTG TGTTAAGGGG TTCAAAAAAT TG 472

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 291 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGCACGAGA TGA ACTCCAA GGAGAAAAG GACCTGGGAA CTCTGGGTAT GACGGTCCCC 60  
 CACCCTGCC CTTGTTGGGA TTCATCAAGA GATGTCATTT GCTGATTGTC TAGGGTGTGG 120  
 CTAATGGGAC CTTGTGTCCT ATCCTTGGCA GGCTACGTGC TGGGCATTAC CATGATGGTG 180  
 ATCATCATTG CCATCGGAGC TGGCATCATC TTGGGCTACT CTACAAGAGG TCAGTAGCTT 240  
 CTCTTCTGGG CCCTCTTAGG AGGAGGGGAG GAAGGTACAC AAAGTCAAAC T 291

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCCGACATG TCTGGAGGCT GTTCTGG 28

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:

a!  
cont

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCGGAAGCT TATTAGGCC CAGGAGTCCC GGC

33

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGCCGGGATC CGCCATCATG CTGTTGGCCT GGGTAC

36

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGCCGGGTAC CTTATTAGGC CCCAGGAGTC CCGGC

35

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

a1  
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGCCGGGATC CGCCATCATG CTGTTGGCCT GGGTAC

36

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

*a1*  
*cont.*  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:



GGCCGGGTAC CTTATTAGGC CCCAGGAGTC CCGGC

35

---

**NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES**

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 C.F.R. 1.821 - 1.825 for the following reason(s):

-  
- ☐ 1. This application clearly fails to comply with the requirements of 37 C.F.R. 1.821-1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
  - ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 C.F.R. 1.821(c).
  - ☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 C.F.R. 1.821(e).
  - ☒ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 C.F.R. 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing."
  - ☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A Substitute computer readable form must be submitted as required by 37 C.F.R. 1.825(d).
  - ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 C.F.R. 1.821(e).
  - ☐ 7. Other: \_\_\_\_\_

**Applicant Must Provide:**

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing".
- ☐ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification.
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 C.F.R. 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d).

For questions regarding compliance to these requirements, please contact:

For Rules Interpretation, call (703) 308-4216

For CRF Submission Help, call (703) 308-4212

PatentIn Software Program Support (SIRA)

Technical Assistance.....703-287-0200

To Purchase PatentIn Software.....703-306-2600

**PLEASE RETURN A COPY OF THIS NOTICE WITH YOUR RESPONSE**